Subject: Re: Post-stratification for DHS data
Posted by Bridgette-DHS on Thu, 17 Nov 2016 17:19:24 GMT
View Forum Message <> Reply to Message

Following is a response from Senior DHS Stata Specialist, Tom Pullum:

Quote:I'm sorry, but I don't use SPSS, and I can't figure out what you did.

I would calculate the total number of weighted cases that you want to have in each district. I think of these as "target" totals. They are obtained by multiplying the total number of cases (unweighted) in your file by the proportions that are in each district in the census. You then multiply the weight (v005) in each district by the ratio of the target total to the current weighted number of cases in the district.

I will illustrate how you would do this in Stata to obtain a uniform distribution of weighted cases across the seven region (v024):

```
set more off
use e:\DHS\DHS_data\KR_files\BDKR70FL.dta, clear
sort v024
save e:\DHS\DHS_data\scratch\BDtemp.dta, replace

keep v024 v005

gen wtd_n=v005/1000000
gen unwtd_n=1
collapse (sum) *wtd_n, by(v024)
list, table clean

* adjust v005 so that the total weight will be the same in each region

summarize unwtd_n
scalar stotal=r(sum)
scalar list stotal
gen target=stotal/7
gen v005_factor=target/wtd_n
list, table clean
keep v005_factor v024

sort v024
merge v024 using e:\DHS\DHS_data\scratch\BDtemp.dta
tab _merge

gen v005_rewtd=round(v005*v005_factor)


*check that the new distribution matches the targets
```

```
gen wtd_n=v005/1000000
gen rewtd_n=v005_rewtd/1000000
gen unwtd_n=1
collapse (sum) *wtd_n, by(v024)
list, table clean
```